MET O 11 TECHNICAL NOTE NO 97

# A Matrix Rank-finding Technique

R. DIXON

## 1. Introduction

Problems and procedures in numerical Meteorology can often be put into matrix-vector form. Once cast in this style the problem can be treated by the powerful techniques of matrix formalism and may often be penetrated to an extent which is difficult by other means. In many cases the validity or effectiveness of a computational sequence depends on the rank of a matrix. This note comments on the question of rank as it relates to computation and also presents a rank-finding algorithm by N.N. Gupta which deserves to be better known.

Met.O.11 (Forecasting Research)
Meteorological Office,
London Road
Bracknell, Berks. U.K.

N.B. This paper has not been published. Permission to quote from it must be obtained from the Assistant Director of the above Meteorological Office Branch.

## 2. The Problem

A matrix A, rectangular or square, is of rank r if all minors of order r + 1 are zero while at least one minor of order r is not zero. A minor of order r is obtained from the parent matrix A by striking out quantum sufficit rows and columns to leave an r-square matrix. The determinant of this r-square matrix is the minor.

This means that the rank of a matrix depends directly upon the values of the determinants of the various sub-matrices which can be formed by the deletion of rows and columns. But the value of a determinant is notoriously liable to be very sensitive to the actual values of the component elements of the determinant. For example

$$\begin{vmatrix} -73 & 78 & 24 \\ 92 & 66 & 25 \\ -80 & 37 & 10 \end{vmatrix} = +1 \qquad (1)$$

but

$$\begin{vmatrix} -73 & 78 & 24 \\ 92 & 66 & 25 \\ -80 & 37 & 10.01 \end{vmatrix} = -118.94 \qquad (2)$$

two orders of magnitude larger and of opposite sign. Furthermore, inspection of (1) and (2) suggests that somewhere between 10 and 10.01 there is a number which would make this determinant zero. There is, but the best I can do with an HP 25 is to say that it lies somewhere between 10.00008337 and 10.00008338.

Another illustration is afforded by considering say the determinant

$$\begin{vmatrix} -73.1234 & -73.1243 & -24.0000 \\ 92.1234 & 92.1243 & 25.0000 \\ -80.1234 & -80.1243 & 100.0000 \end{vmatrix} = 2.1267 \qquad (3)$$

As it stands any matrix having the matrix of (3) as a sub-matrix would be of at least rank 3. However, it is seen that if the numbers in (3) could be entered only to two decimal places the determinant would be zero, since the first two columns would be identical. Thus insofar as it depended upon this particular sub-matrix the parent matrix would not be of rank 3.

So, if a matrix is being discussed mathematically the statement that it is of rank r is precise, but if it is to be used in a computer the statement should rather be that it is of rank r with wordlength m.

## 3. The Rank-finding Method

For any square matrix A the similarity transformation

$$A = R \cdot B \cdot R^{-1} \tag{4}$$

may exist. If A is symmetric it does exist, and furthermore it is possible to find an orthogonal matrix T, remembering that in this case $T^{-1} = \tilde{T}$, such that

$$A = T \cdot D \cdot \tilde{T} \tag{5}$$

where D is the diagonal matrix of eigenvalues of A. From (5), because

$$\tilde{T} \cdot T = I \tag{6}$$

it follows that

$$\begin{aligned} A^2 &= T \cdot D^2 \cdot \tilde{T} \\ A^3 &= T \cdot D^3 \cdot \tilde{T} \\ &\text{- - - - - - - -} \\ &\text{- - - - - - - - -} \end{aligned} \tag{7}$$

etc.

Also needed is the fact that

$$Tr\, A = Tr\,(T \cdot D \cdot \tilde{T}) = Tr\, D \tag{8}$$

where TrA means the trace of A, the sum of the diagonal elements of A.

Even if A is not square and symmetric, $\tilde{A}.A$ or $A.\tilde{A}$ will be. So consider the case where A is a p x n matrix, ($p > n$) of rank r. The maximum rank which a rectangular matrix like this can have is r = n, the dimension of the shortest side, in which case it is said to be of full rank, but for the purpose of this discussion we will assume that A is not of full rank so that $r < n$. We now have

$$\underset{n\rho}{\tilde{A}} \cdot \underset{\rho n}{A} = \underset{nn}{T} \cdot \underset{nn}{D} \cdot \underset{nn}{\tilde{T}} \tag{9}$$

where

$$D = Diag(\lambda_1, \lambda_2, \text{- - -}, \lambda_r, \text{- - -}, \lambda_n) \tag{10}$$

the $\lambda_i$ being the eigenvalues of $\tilde{A}.A$. But, of course

$$\lambda_{r+1} = \lambda_{r+2} = \text{- - - -} = \lambda_n = 0 \tag{11}$$

Consider now the trace of the matrix expression $(I - \alpha \tilde{A} \cdot A)^2$. We have

$$Tr\,(I - \alpha \tilde{A} \cdot A)^2 = Tr\left[(I - \alpha \tilde{A} \cdot A) \cdot (I - \alpha \tilde{A} \cdot A)\right] \tag{12}$$

which by (9) is

$$Tr\,(I - \alpha \tilde{A} \cdot A)^2 = Tr\left[(I - \alpha T \cdot D \cdot \tilde{T}) \cdot (I - \alpha T \cdot D \cdot \tilde{T})\right] \tag{13}$$

i.e.

$$Tr\left(I-\alpha\tilde{A}\cdot A\right)^2 = Tr\left[I - 2\alpha T\cdot D\cdot\tilde{T} + \alpha^2 T\cdot D^2\cdot\tilde{T}\right] \qquad (14)$$

and, using (8), this is

$$Tr\left(I-\alpha\tilde{A}\cdot A\right)^2 = Tr\,I - 2\alpha\, Tr\,D + \alpha^2\, Tr\,D^2 \qquad (15)$$

Evaluating the traces on the RHS, taking note of (10) and (11), we have

$$Tr\left(I-\alpha\tilde{A}\cdot A\right)^2 = n - 2\alpha\left(\lambda_1+\lambda_2+\cdots+\lambda_r\right)+\alpha^2\left(\lambda_1^2+\lambda_2^2+\cdots+\lambda_r^2\right) \qquad (16)$$

But appropriately adding and subtracting r on the RHS this comes to

$$Tr\left(I-\alpha\tilde{A}\cdot A\right)^2 = n-r+\left(1-\alpha\lambda_1\right)^2+\left(1-\alpha\lambda_2\right)^2+\cdots+\left(1-\alpha\lambda_r\right)^2 \qquad (17)$$

Inspection of the algebra from (12) to (17) reveals that evaluating $Tr\left(I-\alpha\tilde{A}\cdot A\right)^4$ would lead to

$$Tr\left(I-\alpha\tilde{A}\cdot A\right)^4 = n-r+\left(1-\alpha\lambda_1\right)^4+\left(1-\alpha\lambda_2\right)^4+\cdots+\left(1-\alpha\lambda_r\right)^4 \qquad (18)$$

and so on. The general result is that

$$Tr\left(I-\alpha\tilde{A}\cdot A\right)^{2k} = n-r+\sum_{i=1}^{i=r}\left(1-\alpha\lambda_i\right)^{2k} \qquad (19)$$

This means that if $\alpha$ is chosen to be $0<\alpha<2/\lambda_1$ the sequence (19) for k = 1, 2, 3, - - -, converges to the nullity (n-r) of A.

It may seem that to use this result it is first necessary to find the largest eigenvalue $\lambda_1$, but this is avoided by using Gerschgorin's theorem which states that $\lambda_1$ is less than or equal to the maximum row sum of magnitudes of the elements of $\tilde{A}.A$. This enables a sufficiently good choice for $\alpha$ to be made.

## 3. Practical Examples

The algebra of Section 2 was translated into a Fortran algorithm "Subroutine Ranker" by Mrs Anne Jackson. The listing is given in Appendix 1. The Subroutine was applied to two matrices. Matrix 1, given in Appendix 2, was an 11 x 50 weighting matrix used by Met.O.19 in connection with the inversion of satellite radiance data. Matrix 2, given in Appendix 3, was a 7 x 7 matrix taken from a book by Cherkasova. The results are shown in Tables 1 and 2.

Table 1 - Matrix 1

| Iteration k | Trace |
|---|---|
| 1 | 9.76 |
| 2 | 9.19 |
| 3 | 8.57 |
| 4 | 7.88 |
| 5 | 7.10 |
| 6 | 6.25 |
| 7 | 5.40 |
| 8 | 4.55 |
| 9 | 3.75 |
| 10 | 3.02 |
| 11 | 2.30 |
| 12 | 1.64 |
| 13 | 1.13 |
| 14 | 0.82 |
| 15 | 0.62 |
| 16 | 0.36 |
| 17 | 0.13 |
| 18 | 0.02 |
| 19 | 0.00 |
| 20 | 0.00 |
| 21 | 0.00 |
| 22 | 0.00 |
| 23 | 0.00 |

Table 2 - Matrix 2

| Iteration k | Trace |
|---|---|
| 1 | 5.16 |
| 2 | 4.41 |
| 3 | 3.56 |
| 4 | 2.71 |
| 5 | 2.05 |
| 6 | 1.66 |
| 7 | 1.40 |
| 8 | 1.16 |
| 9 | 1.02 |
| 10 | 1.00 |
| 11 | 1.00 |
| 12 | 1.00 |
| 13 | 0.99 |
| 14 | 0.99 |
| 15 | 0.97 |
| 16 | 0.95 |
| 17 | 0.90 |
| 18 | 0.81 |
| 19 | 0.65 |
| 20 | 0.43 |
| 21 | 0.18 |
| 22 | 0.03 |
| 23 | 0.00 |

As Matrix 1 is a matrix which has been used successfully, in the purely computational sense, in the inversion of satellite data it was a foregone conclusion that this matrix would be found to be of full rank, in this case rank 11. Table 1 and Figure 1 confirm this, showing the successive traces descending steadily and unequivocably to zero, indicating that Matrix 1 is of nullity zero.

With Matrix 2 the picture is somewhat different. The traces converge eventually to zero but the process clearly has difficulty in getting past the value 1. The possibility exists that the renewed convergence after iteration 12 may be a consequence of round-off effects in the algorithm itself. Even if this possibility can be ruled out the figures in Table 2 and the Figure 2 graph give a strong indication that, with single-length working in the 360/195, Matrix 2 is nearly of nullity 1, i.e. it is nearly of rank 6.

R DIXON
Met.O.11
September 1977

Reference

N.N. GUPTA:- "An optimum Iterative Method for the Computation of Matrix Rank" IEEE Transactions on Systems, Man, and Cybernetics, July 1972.

# APPENDIX 1

FORTRAN IV G LEVEL 21 RANKER

```
0001            SUBROUTINE RANKER(A,M,N)
              C THIS ROUTINE FINDS THE RANK OF THE M*N MATRIX A
0002            DIMENSION A(M,N),B(15,15),C(15,15)
0003            MAXMN=MAXO(M,N)
0004            MINMN=MING(M,N)
0005          2 FORMAT('0'/'0')
0006          3 FORMAT('0',8F12.3)
0007            WRITE(6,2)
0008            DO 130 I=1,M
0009            WRITE(6,3)(A(I,J),J=1,N)
0010        130 CONTINUE
0011            DO 10 I=1,MINMN
0012            DO 10 J=1,MINMN
0013            D=0
0014            IF(MAXMN.EQ.M)GOTO 110
0015            DO 20 K=1,MAXMN
0016            D=D+A(I,K)*A(J,K)
0017         20 CONTINUE
0018            GOTO 120
0019        110 DO 25 K=1,MAXMN
0020            D=D+A(K,I)*A(K,J)
0021         25 CONTINUE
0022        120 B(I,J)=D
0023         10 CONTINUE
0024            WRITE(6,2)
0025            DO 140 I=1,MINMN
0026        140 WRITE(6,3)(B(I,J),J=1,MINMN)
0027            E=0.0
0028            DO 30 I=1,MINMN
0029            D=0.0
0030            DO 35 J=1,MINMN
0031            D=D+ABS(B(I,J))
0032         35 CONTINUE
0033            IF(D.GT.E)E=D
0034         30 CONTINUE
0035            WRITE(6,2)
0036            WRITE(6,3)E
0037            S=1.0/E
0038            WRITE(6,2)
0039            WRITE(6,3)S
0040            DO 40 I=1,MINMN
0041            DO 40 J=1,MINMN
0042            B(I,J)=-S*B(I,J)
0043            IF(I.EQ.J)B(I,J)=1.0+B(I,J)
0044         40 CONTINUE
0045            WRITE(6,2)
0046            DO 150 I=1,MINMN
0047        150 WRITE(6,3)(B(I,J),J=1,MINMN)
0048            D=0.0
0049            DO 50 I=1,MINMN
0050            D=D+B(I,I)
0051         50 CONTINUE
0052            WRITE(6,1)D
0053          1 FORMAT(' TRACE=',F6.2)
0054            DO 100 IJ=1,100
0055            DO 60 I=1,MINMN
```

```
0056              DO 60 J=1,MINMN
0057              D=0.0
0058              DO 70 K=1,MINMN
0059              D=D+B(I,K)*B(K,J)
0060           70 CONTINUE
0061              C(I,J)=D
0062           60 CONTINUE
0063              WRITE(6,2)
0064              DO 160 I=1,MINMN
0065          160 WRITE(6,3)(C(I,J),J=1,MINMN)
0066              D=0.0
0067              DO 80 I=1,MINMN
0068              D=D+C(I,I)
0069           80 CONTINUE
0070              WRITE(6,1)D
0071              DO 90 I=1,MINMN
0072              DO 90 J=1,MINMN
0073              B(I,J)=C(I,J)
0074           90 CONTINUE
0075          100 CONTINUE
0076              RETURN
0077              END
```

# APPENDIX 2 – MATRIX 1

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 79.000 | 119.000 | 168.000 | |
| 217.000 | 296.000 | 405.000 | 632.000 | 860.000 | 1107.000 | 1294.000 | 1531.000 | |
| 1838.000 | 2124.000 | 2391.000 | 2697.000 | 2974.000 | 3290.000 | 3547.000 | 3853.000 | 1st row |
| 4159.000 | 4446.000 | 4604.000 | 4742.000 | 4851.000 | 4940.000 | 4959.000 | 4910.000 | |
| 4673.000 | 4406.000 | 4149.000 | 3823.000 | 3458.000 | 3053.000 | 2579.000 | 2144.000 | |
| 1867.000 | 2816.000 | | | | | | | |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | 0.0 | 0.0 | 89.000 | 178.000 | 266.000 | 400.000 | |
| 533.000 | 710.000 | 888.000 | 1199.000 | 15090.000 | 18650.000 | 23530.000 | 28860.000 | |
| 34180.000 | 39960.000 | 46610.000 | 51940.000 | 55940.000 | 59930.000 | 6038.000 | 5949.000 | 2nd row |
| 5461.000 | 4883.000 | 4351.000 | 3907.000 | 3507.000 | 3152.000 | 2841.000 | 2575.000 | |
| 2353.000 | 2131.000 | 1909.000 | 1731.000 | 1554.000 | 1421.000 | 1243.000 | 1110.000 | |
| 977.000 | 844.000 | | | | | | | |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 90.000 | 288.000 | 587.000 | 826.000 | 1094.000 | 1323.000 | 1602.000 | |
| 1989.000 | 2328.000 | 2765.000 | 3193.000 | 3780.000 | 4367.000 | 4884.000 | 5372.000 | |
| 5750.000 | 6038.000 | 5939.000 | 5759.000 | 5491.000 | 5133.000 | 4675.000 | 4118.000 | 3rd row |
| 3690.000 | 3183.000 | 2716.000 | 2298.000 | 1999.000 | 1781.000 | 1562.000 | 1313.000 | |
| 1104.000 | 925.000 | 686.000 | 507.000 | 398.000 | 239.000 | 129.000 | 60.000 | |
| 20.000 | 0.0 | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 220.000 | 650.000 | 1200.000 | 2180.000 | 3380.000 | 4360.000 | 5010.000 | 5560.000 | |
| 5880.000 | 6100.000 | 5990.000 | 5770.000 | 5450.000 | 5120.000 | 4790.000 | 4360.000 | |
| 3920.000 | 3600.000 | 3380.000 | 3050.000 | 2830.000 | 2610.000 | 2400.000 | 2180.000 | 4th row |
| 1960.000 | 1740.000 | 1420.000 | 1200.000 | 980.000 | 760.000 | 550.000 | 440.000 | |
| 330.000 | 250.000 | 190.000 | 150.000 | 70.000 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | | | | | | | |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 110.000 | |
| 650.000 | 1510.000 | 2910.000 | 4420.000 | 5610.000 | 6690.000 | 7340.000 | 7770.000 | |
| 7980.000 | 7660.000 | 6900.000 | 5830.000 | 5070.000 | 4420.000 | 3780.000 | 3240.000 | 5th row |
| 2700.000 | 2370.000 | 2160.000 | 1830.000 | 1620.000 | 1400.000 | 1190.000 | 970.000 | |
| 860.000 | 760.000 | 650.000 | 540.000 | 430.000 | 320.000 | 220.000 | 110.000 | |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | | | | | | | |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 750.000 | 1930.000 | 3530.000 | |
| 5460.000 | 7070.000 | 8240.000 | 8780.000 | 8890.000 | 8240.000 | 7390.000 | 6420.000 | |
| 5570.000 | 4820.000 | 4180.000 | 3530.000 | 3000.000 | 2460.000 | 2030.000 | 1710.000 | 6th row |
| 1390.000 | 1180.000 | 960.000 | 750.000 | 640.000 | 430.000 | 320.000 | 210.000 | |
| 140.000 | 80.000 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0 | 0.0 | 710.000 | 2750.000 | 5910.000 | 7950.000 | 9070.000 | |
| 9380.000 | 9070.000 | 8260.000 | 7030.000 | 6010.000 | 5200.000 | 4490.000 | 3870.000 | |
| 3360.000 | 2750.000 | 2350.000 | 1940.000 | 1630.000 | 1430.000 | 1220.000 | 1020.000 | |
| 920.000 | 820.000 | 710.000 | 610.000 | 510.000 | 410.000 | 310.000 | 200.000 | 7th row |
| 100.000 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | | | | | | | |
| 2860.000 | 7300.000 | 10370.000 | 12590.000 | 12800.000 | 11110.000 | 8360.000 | 6350.000 | |
| 5080.000 | 4230.000 | 3490.000 | 2750.000 | 2120.000 | 1800.000 | 1480.000 | 1270.000 | |
| 1060.000 | 850.000 | 740.000 | 640.000 | 560.000 | 500.000 | 420.000 | 350.000 | 8th row |
| 290.000 | 240.000 | 180.000 | 130.000 | 90.000 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | | | | | | | |
| 12990.000 | 17380.000 | 17670.000 | 12610.000 | 8790.000 | 6460.000 | 5060.000 | 3920.000 | |
| 2870.000 | 2200.000 | 1820.000 | 1430.000 | 1150.000 | 960.000 | 760.000 | 670.000 | |
| 570.000 | 500.000 | 460.000 | 400.000 | 360.000 | 290.000 | 210.000 | 170.000 | |
| 140.000 | 110.000 | 50.000 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 9th row |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 0.0 | 0.0 | | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 23720.000 | 26020.000 | 14480.000 | 10280.000 | 7140.000 | 5040.000 | 3570.000 | 2410.000 |
| 1470.000 | 1050.000 | 840.000 | 630.000 | 530.000 | 440.000 | 400.000 | 350.000 |
| 320.000 | 300.000 | 240.000 | 210.000 | 180.000 | 150.000 | 120.000 | 100.000 |
| 70.000 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | | | | | | |
| 71700.000 | 17500.000 | 6130.000 | 2830.000 | 1420.000 | 470.000 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | | | | | | |

10th row

11th row

APPENDIX 3 - MATRIX 2

| | | | | | | |
|---|---|---|---|---|---|---|
| 22.1 | - 24.4 | - 11.3 | 18.4 | 15.3 | 20.1 | 25.2 |
| - 24.4 | 16.4 | 16.2 | 14.3 | 10.1 | - 19.5 | 18.8 |
| - 11.3 | 16.2 | 26.1 | - 29.3 | 27.4 | 16.3 | 24.4 |
| 18.4 | 14.3 | - 29.3 | 19.2 | 29.1 | 31.5 | 11.8 |
| 15.3 | 10.1 | 27.4 | 29.1 | 14.9 | 14.9 | 11.3 |
| 20.1 | - 19.5 | 16.3 | 31.5 | 14.9 | 16.8 | 20.8 |
| 25.2 | 18.8 | 24.4 | 11.8 | 11.3 | 20.8 | 18.8 |